DOCUMENT RESUME

ED 152 803 95 TM 006 963

AUTHOR Hahn, Christine T.
TITLE Domain Referenced Testing. An Annotated ERIC Bibliography.
INSTITUTION ERIC Clearinghouse on Tests, Measurement, and Evaluation, Princeton, N.J.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Jul 77
NOTE 22p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Achievement Tests; *Annotated Bibliographies; *Criterion Referenced Tests; Educational Objectives; Elementary Secondary Education; Higher Education; Item Analysis; *Literature Reviews; Norm Referenced Tests; *Test Construction; Test Reviews
IDENTIFIERS *Domain Referenced Tests

ABSTRACT

This 32-item bibliography was compiled to provide access to research and discussions of domain referenced testing. It is not limited to any educational level, nor is it confined to any specific curriculum area. Five data bases were searched by computer. A computer search of the Educational Resources Information Center (ERIC) data base yielded documents announced in Resources in Education and journal articles indexed in Current Index to Journals in Education. Also searched by computer were Psychological Abstracts, Exceptional Child Education Abstracts, Sociological Abstracts, and Comprehensive Dissertation Abstracts. A subject index is provided. (Author/HV)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED152803

TM006 963

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ERIC

# DOMAIN REFERENCED TESTING

# DOMAIN REFERENCED TESTING

# AN ANNOTATED ERIC BIBLIOGRAPHY

Compiled by

Christine T. Hahn

July 1977

ERIC CLEARINGHOUSE ON TESTS, MEASUREMENT, AND EVALUATION
EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY 08540
609-921-9000

## PREFACE

The Educational Resources Information Center (ERIC) is operated by the National Institute of Education of the United States Department of Health, Education, and Welfare. It is an information system dedicated to the improvement of education through the dissemination of conference proceedings, instructional programs, manuals, position papers, program descriptions, research and technical reports, literature reviews, and other types of material. ERIC aids school administrators, teachers, researchers, information specialists, professional organizations, students, and others in locating and using information which was previously unpublished or which would not be widely disseminated otherwise.

The ERIC Clearinghouse on Tests, Measurement, and Evaluation (ERIC/TM) acquires and processes documents and journal articles within the scope of interest of the Clearinghouse for announcement in ERIC's index and abstract journals: Resources in Education (RIE) and Current Index to Journals in Education (CIJE).

Besides processing documents and journal articles, the Clearinghouse has another major function: information analysis and synthesis. The Clearinghouse prepares bibliographies, literature reviews, state-of-the-art papers, and other interpretive reports on topics in its area of interest.

## ABOUT THE BIBLIOGRAPHY

This bibliography was compiled to provide teachers, researchers, and evaluators of educational achievement tests access to information found in journal articles, research papers, books and dissertations concerning domain referenced testing. The primary purpose of these tests is to estimate the extent to which a student has attained or retained the intended learning outcomes of a particular segment of instruction. Domain referenced testing (DRT) is particularly useful in ascertaining the learner's strengths and weaknesses in a specific subject area. This bibliography is not limited to any educational level, nor confined to any specific curriculum area. Five data bases were searched by computer for this bibliography.

ERIC data base yielded documents announced in Resources in Education and journal articles indexed in Current Index to Journals in Education which covers over 700 education-related journals. Psychological Abstracts, an index providing summaries of literature in psychology and related disciplines, covers over 800 journals, technical reports, monographs, and other scientific documents. Exceptional Child Education Abstracts (CEC), a data base concerned with published and unpublished literature on the education of handicapped and gifted children, covers such sources as books, journal articles, teaching materials, and reports. Sociological Abstracts, an index covering literature in sociology and related disciplines, scans over 1200 journals and serial publications a year. Comprehensive Dissertation Abstracts is a definitive subject, title, and author guide to virtually every American dissertation accepted at

an accredited institution since 1861, and to thousands of Canadian dissertations.

For ERIC documents (those with an ED number appearing at the end of the bibliographic citation) the following information is presented when available: Personal or corporate author, title, date of publication, number of pages, and availability information. These documents may be purchased in hard copy or in microfiche from the ERIC Document Reproduction Service (EDRS). Price information and an order form are appended. However, ERIC microfiche collections are available at approximately 590 locations throughout the country, and most of these collections are open to the public. If you are unable to find a collection in your area, you may write ERIC/TM for a listing. Documents with a UMI Order number can be obtained from: University Microfilms International, P.O. Box 1764, Ann Arbor, Michigan 48106.

Journal articles (those entries appearing with an EJ number or otherwise identified as journals by the bibliographic citation) are not available from EDRS. However, most of these journals are readily available in college and university libraries as well as some large public libraries.

All entries are listed alphabetically by author and are numbered. An abstract, or in the case of most journal articles, a shorter annotation, is provided for each entry. A subject index consisting of ERIC descriptors and identifiers reflecting major emphasis is also provided. Numbers appearing in the index refer to entries.

1. Baker, Eva L. Beyond Objectives: Domain-Referenced Tests for Evaluation and Instructional Improvement. Educational Technology, Vol. 14, No. 6, June 1974, pages 10-16.

   This article describes the inadequacy of current methods of assessing attainment of behavioral objectives. It is suggested that domain-referenced testing be more frequently utilized. Guidelines for preparing domains are presented, with illustrative examples.

2. Baker, Eva L. Using Measurement to Improve Instruction. September 1972. 8 pages. ED 069 762.

   Instructional improvement within the context of criterion-referenced and norm-referenced tests is described. Such categories overemphasize test interpretation rather than design characteristics of achievement tests. Data from most measurement situations may be reported or interpreted either according to criterion- or norm-referenced standards. How the test is developed and what it represents is of critical importance. The paper proposes alternative conceptualizations of test design: construct-referenced, objectives-referenced, and domain-referenced. Using student data, the teacher needs to identify deficiencies in achievement, possible explanations, and remedies, and to put the remedies into operation. An analysis of the utility of each test type results in the appraisal that domain referenced tests provide the most information for teachers and therefore are the most desirable as data sources for instructional improvement. However, because of lack of knowledge about instruction, poor training in available instructional principles, and lack of resources to encourage changes in instructional habits, it is concluded that instructional improvement, even if measurement considerations were satisfied, is not imminent.

3. Besel, Ronald, and Okada, Masahito. The Development of Domain-Referenced Tests for an Objectives-Based Reading Program. April 1974. 8 pages. ED 093 918.

   Criteria for the selection of item forms, content domains, and sampling procedures for program specific, domain-referenced tests are developed. The primary purpose of these tests is to estimate the extent to which individual pupils have attained or retained the intended learning outcomes of a particular segment of instruction. Tests developed for the tryout of the SWRL Reading Program illustrate the application of the criteria. A variety of critical reading skills is assessed. The use and potential value of facet designed tests for assessing word recognition and novel word decoding is described. Error type scores provide potentially valuable information on which to base prescriptions of supplementary instruction.

4. Denham, Carolyn H. Criterion-Referenced, Domain-Referenced and Norm-Referenced Measurement: A Parallax View. Educational Technology, Vol. 15, No. 12, December 1975, pages 9-12.

Three measurement techniques, criterion-referenced (CR), domain-referenced (DR), and norm-referenced (NR) tests are defined, analyzed, and compared in this article. Explained and underscored is the clear separation of criterion-referenced tests from domain-referenced tests, a separation seldom made by education experts. Although both tests utilize a random sample of items drawn from a larger domain of items, criterion-referenced tests compare the examinee to a specific criterion, while domain-referenced tests are more concerned with ascertaining the examinee's individual strengths and weaknesses in a particular subject area. Also, the value of combining the three types of tests (CR, DR, and NR) to construct new tests is suggested. Finally, the need for reliable and valid item selection procedures is stated. Possible selection processes for CR, DR, and NR tests are outlined, as well as for the four combination-type tests suggested by the author.

5. Duncan, Ann Dell. Tracking Behavioral Growth: Day-to-Day Measures of Frequency Over Domains of Performance. Educational Technology, Vol. 14, No. 6, June 1974, pages 54-59.

In this article ways in which domain-referenced testing may be utilized in a program of facilitating personal growth are described. The personal growth program described utilizes behavioral principles in helping participants to change. Its use together with the innovative testing approach is discussed in terms of implementation and advantages.

6. Durnin, John, and Scandura, Joseph M. An Algorithmic Approach to Assessing Behavior Potential: Comparison with Item Forms and Hierarchical Technologies. Journal of Educational Psychology, Vol. 65, No. 2, October 1973, pages 262-272.

In this study two item form technologies, the item forms technology (domain-referenced testing) of Hively (1968), and the hierarchical or stratified forms technology of Ferguson (1969), were compared with an algorithm-based technology for assessing behavior potential. Bases for comparison were (a) relative effectiveness in predicting performance on individual test items, based on performance on items identified according to respective technologies; (b) relative power (generalizability); (c) relative efficiency (number of items); and (d) relative validity of item hierarchies. Two parallel tests on column subtraction were administered to 25 subjects. Test performance was analyzed according to each technology. Algorithmic technology (a) better predicted individual subjects failure on individual second test items, (b) had higher generalizability levels, (c) was more efficient, and (d) had higher validity indices on hierarchical ordering of tasks than item form technologies. Implications for diagnostic testing and remediation were discussed.

7. Ferguson, Richard L., and Tse-Chi, Hsu. The Application of Item Generators for Individualizing Mathematics Testing and Instruction. May 1971. 21 pages. ED 053 935.

   Described is a procedure for utilizing a computer to generate domain-referenced tests in mathematics. The procedure can be adapted for use in testing and instructional programs in either an on-line or off-line mode. It requires specification of the objectives of interest in behavioral terms and grouping them into sets that share a common content. Addition, multiplication, and fractions are examples of possible groupings. To implement the procedure, one of the sets of objectives resulting from the grouping process is selected, and item forms representative of the behaviors implied by each objective in the set are specified. Then an item generator is developed that facilitates the construction of items representative of all item forms so identified. Given an on-line computer capability, the authors describe how it is possible to use the proposed item generator for assisting measurement and instruction in an individualized mathematics program.

8. Geisinger, Kurt F. A Systems Approach to Item Production and Review in a Computer Managed Instruction Project. April 1976. 20 pages. ED 121 280.

   An item generation procedure is described which was utilized in the development of Computer Managed Review and Examination courses for the education of nurses in remote areas. The major emphases are the processes of domain definition, item writing, and item edition. Specific discussion is presented concerning methods of item construction to assess technical vocabulary, concept learning, and the application of nursing principles to the solution of problems. The entire test construction procedure is briefly reviewed; this procedure includes numerous quality checks to insure the production of both high caliber instructional materials and domain-referenced tests. The criteria used at various editing and review stages are mentioned. An initial evaluation of the items is made, and problems inherent in the item generation procedure are offered.

9. Haladyna, Thomas. An Analysis of Two Procedures for Decision Making When Using Domain-Referenced Tests. April 1975. 22 pages. ED 104 957.

   A central problem for the user of domain-referenced tests in instruction is deciding who has passed and who has failed. Two procedures were presented and discussed. The first, employing classical test theory, was found to be more useful for larger domains and where the passing standard is 70 percent or less. The sampling procedure suggested by Millman (1974) was found to be more applicable when the test size approximates the size of the domain. Neither procedure appears useful when the passing standard is high. In light of the large numbers of examinees classified as uncertain when real test data is used, it was concluded that neither procedure offers much to decision making in systematic individualized instruction.

10. Haladyna, Thomas. The Paradox of Criterion-Referenced Measurement. April 1976. 25 pages. ED 126 155.

The existence of criterion-referenced (CR) measurement is questioned in this paper. Despite beliefs that differences exist between two alternative forms of measurement, CR and Norm Referenced (NR), an analysis of philosophical and psychological descriptions of measurement, as well as a growing number of empirical studies, reveal that the common distinctions drawn between CR and NR measurement focus on what occurs prior to and following measurement, namely the writing of items and the interpreting of test scores. In this respect, the use of the term "criterion-referenced measurement" is paradoxical. The purpose, method of construction, and usefulness of domain-referenced tests are also discussed in this article, the domain-referenced tests being treated as a particular type of criterion-referenced measurement.

11. Haladyna, Thomas M. The Quality of Domain-Referenced Test Items. April 1976. 28 pages. ED 129 846.

The objectives of this study were to first determine whether or not the empirical item analysis of domain referenced tests (DR) was justified; and second, in the event that it was, which of a set of recommended procedures was most effective for determining item quality. The analysis that followed led to the conclusion that empirical procedures were highly desirable. When these empirical procedures were applied to test data, the results indicated that four different techniques provided almost identical information: Rasch statistics, instructional sensitivity indexes, traditional statistics, and Bayesian indexes. Based on these results, it would seem that any one of these four would serve adequately.

12. Hentschke, Guilbert C., and Levine, Donald M. Planning for Evaluation in Performance Contracting Experiments: The Connection to Domain-Referenced Testing Theory. Educational Technology, Vol. 14, No. 6, pages 38-43.

This article delineates the impact of incorporating domain-referenced testing concepts into performance contracts with teachers. The effect of domain-referenced testing theory on 11 problems resulting from this assessment is described. It is suggested that this approach would alleviate some of these testing difficulties.

13. Hively, Wells. Domain Referenced Testing. October 1974. 150 pages. Available from Educational Technology Publications, Englewood Cliffs, New Jersey 07632 ($4.95).

The central assumption in domain-referenced testing (DRT), as presented in this book, is that a domain may be determined which adequately

represents a particular universe of knowledge. After a domain has been established, the technological and practical problem of using domain-referenced testing must be solved. This book contains a collection of twelve short chapters covering such DRT topics as definition and function; sampling plans; instructional accountability; curriculum assessment, management, and modification; teacher, program, and product evaluation; relation of performance contracting experiments to DRT; individualized instruction; and behavioral growth tracking. Brief comments and helpful sources are provided by the editor.

14. Hively, Wells. Introduction to Domain-Reference Testing. Educational Technology, Vol. 14, No. 6, June 1974, pages 5-10.

This article describes the theory and utilization of the domain-referenced approach to the measurement and technology of educational objectives. According to this method, sample problems are generated in ways clearly specifiable before the test. Thus, a clearly specified domain of competence exists and is available to the test taker prior to the test. Domain-referenced testing has its roots in learning theory and collects data useful in evaluating growth. Its more traditional alternative, norm-referenced testing (NRT), has its roots in the study of individual differences so that the structure of the content is not considered important. NRT collects data useful in prediction and selection but not in evaluating instruction. Education requires both types of testing but the latter has been emphasized traditionally.

15. Hively, Wells, Ed., and Reynolds, Maynard C., Ed. Domain-Referenced Testing in Special Education. 1975. 146 pages. Available from Council for Exceptional Children, 1920 Association Drive, Reston, Virginia 22091 ($4.00), Product 101.

Presented are eight papers that deal with the educational implications for handicapped children of domain-referenced testing, as contrasted with standardized norm-referenced achievement testing. The crucial aspects of each testing model are highlighted by W. Hively in an introductory section. M. Reynolds surveys past and present special education pressures and analyzes their impact on testing. T. Donlon reviews historical and technical concepts of test-score referencing and points out complexities and confusions in terminology among different types of evaluation. Discussed by J. Rosner are test construction and utilization in connection with an adaptive perceptual skills curriculum. Explained by A. Hofmeister are procedures and materials for training teachers to integrate criterion-referenced testing and instruction within the regular classroom. The creation of a comprehensive computer-based information bank in the area of reading instruction and its use in domain-referenced test development is described by R. O'Reilly. Examined is the use of domain-referenced testing in the delivery of special education services in a rural area (F. Hammar-

back and C. Koenig). Ethical considerations in the use of norm-domain-, and behavior-referenced testing are considered in the final paper by E. Joselyn. Also included are a 60-item bibliography on domain-referenced testing and biographical information about the authors.

16. Johnston, Thomas J. Program and Product Evaluation from a Domain-Referenced Viewpoint. Educational Technology, Vol. 14, No. 6, June 1974, pages 43-48.

    Factors to be considered in analyzing educational products and programs within the domain-referenced testing framework are described in this article. This analysis is discussed in terms of the characterization of domains and the application of domain weighting.

17. Kwansa, Kofi Bassa. Investigation of the Relative Content Validity of Norm-Referenced and Domain-Referenced Arithmetic Tests. Ph. D. Dissertation, University of Pittsburgh, 1972. (UMI Order No. 73-4153, 256 pages.)

    Norm-referenced and domain-referenced methods were each used to build sixth grade arithmetic tests. The tests were administered to samples of students and the results used for making content validity comparisons between the tests. Findings showed that the domain-referenced tests had higher content validity than the norm-referenced tests, that parallel forms of the norm-referenced tests did not show equivalent degrees of content validity between themselves, that scores on the norm-referenced tests correlated highly with scores on the domain-referenced tests, and that the domain-referenced tests had slightly smaller standard errors of estimation and prediction than the norm-referenced tests.

18. Macready, George Byron. An Investigation into the Nature of Interitem Relations and the Structure of Domain Hierarchies Found Within a Domain Referenced Testing System. Dissertation Abstracts International, Vol. 33, No. 5-A, 1972, page 2174.

    The purpose of this study was to establish procedural techniques which might be helpful in the assessment of achievement testing systems which use operationally specified procedures for both the generation and grouping of items. In addition, this study attempted to assess the relations among items generated by a Domain Referenced Testing System within the curriculum area of multiplication of whole numbers. It was hoped that such an assessment would provide information on the degree to which it is possible to group such items into sets or "domains" of equivalent items. Such information was of interest when the operation

procedures for grouping items were based on the assumed processes involved in arriving at answers to the items. It was of further interest to determine the order (or partial order) in which the skills necessary for correctly answering items from the various domains were acquired by students. Here an attempt was made to determine both the nature and extent to which such a partial ordering could be established. In general, it was possible to infer from the results of this study that the Domain Referenced Testing System studied provided an effective means of grouping items into sets of "equivalent" items (i.e., items which a given student tended to answer either all correctly or all incorrectly). Thus, this testing system allowed for an accurate description of how students could be expected to perform on an entire domain of items on the basis of a small sample of items.

19. Macready, George B. The Structure of Domain Hierarchies Found Within a Domain Referenced Testing System. Educational and Psychological Measurement, Vol. 35, No. 3, Autumn 1975, pages 583-597.

In this article conditional states of item mastery found among items from different item domains and the effectiveness of various procedures for identifying such conditional relations were assessed. The item domains considered were from the curriculum area of multiplication of whole numbers, and were defined by a domain referenced testing system. Data were gathered during pilot and main studies from a total of 400 5th graders. It was possible to infer from the results of this study that the domain referenced testing system considered produced items which across domains showed strong conditional relations. Comparisons of goodness of fit were made among domain hierarchies with similar numbers of specified conditional relations generated by 2 different empirical procedures and by experts' judgment. Additional comparisons were made among models generated by the same procedure but with different numbers of specified conditional relations. Support for the validity of empirically generated hierarchies with moderate numbers of conditional relations among domains was provided.

20. Macready, George B., and Merwin, Jack C. Homogeneity Within Item Forms in Domain Referenced Testing. Educational and Psychological Measurement, Vol. 33, No. 2, Summer 1973, pages 351-360.

This article studied the nature of the relationships found in domain-referenced tests among items within item forms and how these relationships compare with an ideal case for diagnostic tests in which, if a person gets 1 item within an item form right, then he would get all items within the item form correct. Subjects were 91 corpsmen from 5 randomly chosen Youth Conservation Centers. Each subject was administered a 75-item test on the multiplication of whole numbers which had been generated from 25 item forms based on intuitive categories. Results show that, in most cases, item

forms which generate items of moderate difficulty can be used to obtain relatively homogeneous sets of items of equivalent difficulty for a defined population of subjects. Such item forms provide sets of items superior to those which would be expected if item difficulties alone were used to group items into sets.

21. Millman, Jason. Criterion-Referenced Measurement. In Popham, W.J. Evaluation in Education. Berkeley: McCutchan Pub. Corp., 1974.

This chapter should not only acquaint the reader with the present state of the art on Criterion-Referenced (CR) measurement but also suggest possible directions for further inquiry. The goal of the first part of this chapter is to deal with the definitional dilemma of CR measurement by proceeding from the more traditional view of CR measurement to one that is more productive and provides a unifying theme for the study. The focus of the second part of the chapter is on tests intended to describe the current status of an examinee with respect to a well-explicated set of performance tasks called a domain. A random, or stratified random, sample of items from a domain is called a domain-referenced test (DRT). Specific topics include defining the item population, selecting test items, establishing a passing score, determining test length, and evaluating the DRT. Tests having the function of discriminating between individuals or groups of individuals believed to differ on the attribute purportedly measured by the test are called differential assessment devices (DAD's). Some DAD's reference a particular objective or skill with sufficient specification that a criterion-referenced interpretation is reasonable. The development and evaluation of such tests, labeled CRDAD's is presented in the third section of the chapter. Finally, selected areas of application in education which call for measuring status or differentiating individuals or groups are discussed.

22. Millman, Jason. Sampling Plans for Domain-Referenced Tests. Educational Technology, Vol. 14, No. 6, June 1974, pages 17-21.

A way of assigning items in a domain-referenced testing plan so that examinees encounter them in orders not affecting their subsequent responses is described in this article. A sampling scheme for carrying this out is presented. Ways of using such a scheme and possible sources of bias are also discussed.

23. Nitko, Anthony J. Some Considerations When Using a Domain-Referenced System of Achievement Tests in Instructional Situations. March 2, 1970. 24 pages. ED 037 793.

The problem of using a domain-referenced system of achievement tests is discussed as it relates to the design of instruction. Testing problems are discussed from the point of view that the teacher, pupil, and/or automation needs certain kinds of information in order to make instructional decisions that are adaptive to the individual learner. The design of achievement tests based on item forms is determined by

the purpose(s) for which the information obtained from them is needed. The selection of items from the defined domain of item forms is discussed in terms of the purpose for testing, the relationship between items and instruction, and the relationship between instructional objectives and item forms.

24. Nitko, Anthony J., and Hsu, Tse-Chi. Using Domain-Referenced Tests for Student Placement, Diagnosis, and Attainment in a System of Adaptive, Individualized Instruction. Educational Technology, Vol. 14, No. 6, June 1974, pages 48-54.

    This article illustrates ways in which domain-referenced testing might be used in an adaptive and individualized system of instruction. It is suggested that measurement and instruction should be integrated into a decision-making context. Examples are provided.

25. Olympia, P. L. Jr. Repetitive Domain-Referenced Testing Using Computers: the TITA System. June 1975. 9 pages. ED 111 358.

    The TITA (Totally Interactive Testing and Analysis) System algorithm for the repetitive construction of domain-referenced tests utilizes a compact data bank, is highly portable, is useful in any discipline, requires modest computer hardware, and does not present a security problem. Clusters of related key phrases, statement phrases, and distractors form minipools from which the computer generates items for a domain-referenced unit of instruction. Test items can take the form of multiple-choice, true-false, matching, and fill-in questions. A random number generator produces data for test items requiring numerical solutions, and the correct answer is computed from a coded formula so computational subroutines are not required for each test item. This component of computer managed instruction allows the instructor to key related items in the data minipool to learning resources and to code the resources themselves for inclusion in the data bank. Use of this system for elementary, secondary, or undergraduate courses can facilitate instructional management and result in positive effects on student morale.

26. O'Reilly, Robert P., and others. The Validation and Refinement of Measures of Literal Comprehension in Reading for Use in Policy Research and Classroom Management. February 1976. 424 pages. ED 133 363.

    The report proposes to complete the validation and refinement of a new domain referenced testing technology designed to assess literal comprehension ability in students in grades 1-12. The domain referenced measures in this technology, along with other more traditional measures of reading comprehension, literal and non-literal, are subsequently intended to be used in part in large

scale studies of productivity in school reading programs. To date, studies of productivity in reading instruction have had little influence on educational decision-making due to serious methodological problems, one of the major problems being the lack of adequate measures of program output. The report further proposes to solve a number of important instructional management problems created by the use of the inadequate information available from traditional measures of reading comprehension. The new domain referenced measures of reading comprehension will have an improved basis for scaling students on comprehension ability, and ability scores from this scale will be referenced to an additional scale defining an individual or group's ability to read in several domains of written discourse. These scaling features will allow for the assignment of students to specific levels of reading materials in specific instructional or content domains, a procedure not possible with existing measures of reading comprehension.

27. Popham, W. James. An Evaluation Guidebook; A Set of Practical Guidelines for the Educational Evaluator. 1972. 89 pages. Available from Instructional Objectives Exchange, Box 24095, Los Angeles, California 90024 ($2.50)

The third chapter of this book discusses the topic of domain-referenced testing (DRT) in detail. DRT is seen as a useful measuring device in determining whether or not an educational objective has been accomplished. Its essential ingredient involves defining the domain of learner behaviors called for in the objective, then referencing all test items to this domain. The next procedure in constructing a domain-referenced test is the preparation of an item form which contains three necessary elements: (1) instructions to students; (2) stimulus limits; and (3) response limits. These elements are defined and discussed by the author, and two illustrative item forms are presented.

28. Popham, James W. Teacher Evaluation and Domain-Referenced Measurement. Educational Technology, Vol. 14, No. 6, June 1974, pages 35-37.

The effect of domain-referenced measurement on teacher evaluation is discussed in this article. This approach corrects most of the deficits of standardized tests. Because the domain-referenced approach produces clear categories of learner behaviors to be measured, it enables teachers to know better where their teaching has not worked. Ways of improving teacher performance through use of this method are described.

29. Sanders, James R., and Murray, Stephen L. Alternatives for Achievement Testing. Educational Technology, Vol. 16, No. 3, March 1976, 7 pages.

Explored are four selected strategies--norm-referenced, criterion-referenced, objectives-referenced, and domain-referenced testing--of achievement test development, as well as implications for their application. Each type of testing approach is discussed in terms of such aspects as definition, key emphasis, development procedure, item selection, necessary input for test development, types of scores reported, examples of test interpretation, recommended uses, and inappropriate uses and limitations. It is concluded that the best achievement testing system is probably a combination or variation of the approaches.

30. Sension, Donald B., and Rabehl, George. Test-Item Domains and Instructional Accountability. Educational Technology, Vol. 14, No. 6, June 1974, pages 22-28.

This article discusses the history, goals, and implementation of instructional accountability in education. Domain-referenced testing is suggested as a way of increasing and assessing such accountability. Application of this approach in 2 school systems is described. Evaluation of its success is discussed.

31. Whitely, Susan E. Domain Referenced Testing: An Alternative Model for Test Construction. Proceedings of the Annual Convention of the American Psychological Association, Vol. 6, Pt. 2, 1971, pages 515-516.

Domain-referenced testing, interpreting scores with direct reference to the domain of item content, has been given increasing attention in recent years. Neither the programed learning approach nor the achievement test approach has been able to provide models that can handle complex and heterogeneous domains to allow a domain-referenced score interpretation. A modified version of Stephenson's structured Q-sample model is presented to provide an alternative method of test construction. It is different from current approaches because it provides information concerning domain structure and does not depend upon random sampling to estimate true score.

32. Willoughby, Lee, and others. A Comparison of Domain-Referenced and Classic Psychometric Test Construction Methods. 1976. 13 pages. ED 131 128.

This study compared a domain referenced approach with a traditional psychometric approach in the construction of a test. Results of the December, 1975 Quarterly Profile Exam (QPE) administered to 400 examinees at a university were the source of data. The 400 item QPE is a five alternative multiple choice test of information a "safe" physician should know. Content of the exam covers the broad areas of Internal Medicine, Pediatrics, Obstetrics/Gynecology, Surgery, and Basic Science, as well as additional sub-topics. For purposes of this study, two 75 item tests were constructed by pulling from

the 400 item QPE by two different strategies. The domain referenced approach was used to construct a 75 item test by a random sample of the 400 items. Selection of the 75 items with the highest point biserial item-total correlations represented the traditional psychometric approach to test construction. The exams were then rescored to obtain scores and item analysis data on the random and psychometric tests. Then, the two tests were compared with respect to distribution of p values (the proportion answering an item correctly), point biserial item-total correlations, student scores across medical school year level and reliability. The results were discussed with regard to their consistency with expectations of the domain referenced and psychometric approaches.

## Subject Index

(Major Descriptors and Identifiers Only)


Achievement Tests, 2, 18, 19, 23, 29, 31

Algorithms, 6

Answer Sheets, 27

Behavior, 6

Behavior Change, 13

Behavioral Objectives, 13, 27

Cloze Procedure, 26

Comparative Analysis, 10, 32

Computer Assisted Instruction, 7, 8, 25

Criterion-Referenced Tests, 4, 8, 10, 11, 13, 15, 20, 21, 29, 32

Curriculum Evaluation, 16

Decision Making Skills, 9

Diagnostic Tests, 20, 24

Educational Accountability, 12, 30

Educational Psychology, 6

Educational Trends, 15

Elementary School Mathematics, 17

Evaluation, 12, 13, 23

Evaluation Criteria, 4, 28

Evaluation Methods, 1, 13, 14, 22

General Education, 29

Growth Patterns, 5

Handicapped Children, 15

Homogeneous Grouping, 20

Individualized Instruction, 7, 9

Instructional Improvement, 1, 2

Item Analysis, 6, 11, 18, 20, 30

Item Banks, 8

Item Sampling, 19, 20

Learning Theories, 23

Mathematics Education, 7

Measurement Techniques, 2, 4, 5, 14, 21, 27

Multiple Choice Tests, 26

Multiplication, 19

Norm Referenced Tests, 4, 10, 29, 32

Objectives, 1

Performance Contracts, 12

Performance Criteria, 5

Prediction, 6

Program Evaluation, 16

Programed Instruction, 31

Reading Comprehension, 26

Reading Tests, 3, 26

Research, 17

Sampling, 22

Scoring, 31

Serial Learning, 19

Standardized Tests, 15

State of the Art Reviews, 21

Statistical Analysis, 9, 11, 31

Student Evaluation, 5, 24

Student Placement, 24

Teacher Evaluation, 28

Test Construction, 2, 3, 7, 8, 17, 18, 25, 27, 29, 31, 32

Test Development, 29

Test Interpretation, 2, 9, 23

Test Items, 18, 27

Test Reliability, 21

Test Validity, 14, 20, 22, 26

Testing, 1, 7, 9, 12, 14, 16, 22, 24, 28, 30, 31

Testing Problems, 15, 23

Tests, 4, 1

# ORDER FORM

**ERIC** DOCUMENT REPRODUCTION SERVICE
P.O. Box 190 ARLINGTON, VIRGINIA 22210 • (703) 841-1212

OPERATED BY: COMPUTER MICROFILM INTERNATIONAL, CORP.

SHIP TO: _____

BILL TO: _____

| ED NUMBER | NO. OF PAGES | NO. OF COPIES MF | NO. OF COPIES HC | UNIT PRICE | TOTAL |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| TOTAL NO. OF PAGES | | | | SUBTOTAL | |
| TAX EXEMPT NO. _____ | | VA RESIDENTS ADD 4% SALES TAX | | | |
| DEPOSIT ACCT. NO. _____ | | POSTAGE | | | |
| | | TOTAL | | | |

## IMPORTANT INSTRUCTIONS

- **ORDER BY ED NO.** (3 digits)
  See Resources in Education
- **SPECIFY EITHER:**
  Microfiche (MF)
  or
  Paper Copy (HC)
- **ENTER UNIT PRICE**
  (See Below)
- **INCLUDE POSTAGE**
  (See Below)
- **ENCLOSE CHECK or MONEY ORDER**
  Payable to EDRS or Computer Microfilm International Corp.
  (U.S. Funds Only)
- **MAIL TO:**
  EDRS
  P. O. Box 190
  Arlington, Virginia 22210
- **COMPLETE AND SIGN BELOW**

Date _____

Signature _____

Title _____

### UNIT PRICE SCHEDULE
(Effective June 10, 1976)

| MICROFICHE (MF) | | PAPER COPY (HC) | |
|---|---|---|---|
| NUMBER FICHE EACH ED # | Price | NUMBER PAGES EACH ED # | Price |
| 1 to 5 | $ .83 | 1 to 25 | $1.67 |
| 6 | 1.00 | 26 to 50 | 2.06 |
| 7 | 1.16 | 51 to 75 | 3.50 |
| 8 | 1.33 | 76 to 100 | 4.67 |
| Each additional microfiche | .167* | Each additional 25 pages | 1.34 |

*Total Price Should Be Rounded to Nearest Cent

### CHART FOR DETERMINING UNITED STATES POSTAGE

| 1st CLASS POSTAGE FOR | | 4TH CLASS POSTAGE FOR TOTAL MF OR HC PAGES INDICATED (Allow 3-4 weeks delivery time from date of order) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1-3 Microfiche Only | 4-7 Microfiche Only | 60 or less MF or HC Pages | 61-120 MF or HC Pages | 121-180 MF or HC Pages | 181-240 MF or HC Pages | 241-300 MF or HC Pages | 301-360 MF or HC Pages | 361-420 MF or HC Pages | Each Additional 60 MF or HC Pages |
| .13 | .24 | .25 | .35 | .45 | .55 | .65 | .75 | .85 | .08 |

*Costs For Priority Shipment Available Upon Request.

For Foreign Postage SEE REVERSE—

## GENERAL INFORMATION

**1. PRICE LIST**
The prices set forth herein may be changed without notice; however, any price change will be subject to the approval of the National Institute of Education Contracting Officer.

**2. PAYMENT**
The prices set forth herein do not include any sales, use, excise, or similar taxes which may apply to the sale of microfiche or hard copy to the Customer. The cost of such taxes, if any, shall be borne by the Customer.

Payment shall be made net thirty (30) days from date of invoice. Payment shall be without expense to CMIC.

**3. REPRODUCTION**
Express permission to reproduce a copyrighted document provided hereunder must be obtained in writing from the copyright holder noted on the title page of such copyrighted document.

**4. CONTINGENCIES**
CMIC shall not be liable to Customer or any other person for any failure or delay in the performance of any obligation if such failure of delay (a) is due to events beyond the control of CMIC including, but not limited to, fire, storm, flood, earthquake, explosion, accident, acts of the public enemy, strikes, lockouts, labor disputes, labor shortage, work stoppages, transportation embargoes or delays, failure or shortage of materials, supplies or machinery, acts of God, or acts or regulations or priorities of the federal, state, or local governments; (b) is due to failures of performance of subcontractors beyond CMIC's control and without negligence on the part of CMIC; or (c) is due to erroneous or incomplete information furnished by Customer.

**5. LIABILITY**
CMIC's liability, if any, arising hereunder shall not exceed restitution of charges.

In no event shall CMIC be liable for special, consequential, or liquidated damages arising from the provision of services hereunder.

**6. WARRANTY**
CMIC MAKES NO WARRANTY, EXPRESS OR IMPLIED, AS TO ANY MATTER WHATSOEVER, INCLUDING ANY WARRANTY OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE.

**7. QUALITY**
CMIC will replace products returned because of reproduction defects or incompleteness. The quality of the input document is not the responsibility of CMIC. Best available copy will be supplied.

**8. CHANGES**
No waiver, alteration, or modification of any of the provisions hereof shall be binding unless in writing and signed by an officer of CMIC.

**9. DEFAULT AND WAIVER**
a. If Customer fails with respect to this or any other agreement with CMIC to pay any invoice when due or to accept any shipment as ordered, CMIC may without prejudice to other remedies defer any further shipments until the default is corrected, or cancel this Purchase Order.
b. No course of conduct nor any delay of CMIC in exercising any right hereunder shall waive any rights of CMIC or modify this Agreement.

**10. GOVERNING LAW**
This Agreement shall be construed to be between merchants. Any question concerning its validity, construction, or performance shall be governed by the laws of the State of New York.

**11. DEPOSIT ACCOUNTS**
Customers who have a continuing need for ERIC documents may open a Deposit account by depositing a minimum of $200.00. Once a deposit account is opened, ERIC documents will be sent upon request, and the account charged for the actual cost and postage. A monthly statement of the account will be furnished.

**12. STANDING ORDER ACCOUNTS**
Customers who desire to receive microfiche copies of all ERIC reports announced in each issue of Resources in Education may do so by depositing $2000.00 or submitting an executed purchase order. The cost of each issue and postage will be charged against the account. A monthly statement of the account will be furnished.

**13. PAPER COPY (HC)**
A paper copy (HC) is xerographic reproduction, on paper, of the original document. Each paper copy has a Vellum Bristol cover to identify and protect the document.

**14. FOREIGN POSTAGE**
Postage for all countries other than the United States is based on the international Postal Rates in effect at the time the order is shipped. To determine postage allow 60 microfiche or 60 (HC) pages per pound. Customers must specify the exact classification of mail desired, and include the postage for that classification with their order. Payment must be in United States funds.

## OTHER ERIC COLLECTIONS AVAILABLE FROM EDRS

**STANDING ORDERS**

Subscription orders of microfiche copies of all ERIC reports announced in each issue of **Resources in Education** average $160.00 per month at the rate of 8.7¢ per microfiche. Postage extra.

**BACK COLLECTIONS** (postage extra)

| | |
|---|---|
| Reports in *Research in Education for* 1966 and 1967 | $ 385.06 |
| Reports in *Research in Education for* 1968 | 1,159.36 |
| Reports in *Research in Education for* 1969 | 1,383.21 |
| Reports in *Research in Education for* 1970 | 1,408.36 |
| Reports in *Research in Education for* 1971 | 1,643.69 |
| Reports in *Research in Education for* 1972 | 1,701.28 |
| Reports in *Research in Education for* 1973 | 1,481.70 |
| Reports in *Research in Education for* 1974 | 1,548.60 |
| Reports in *Resources in Education for* 1975 | 1,734.61 |
| Reports in *Resources in Education* Jan.–Apr. 1976 | 545.92 |
| Entire Collection | $12,991.79 |

**AIM/ARM MICROFICHE COLLECTIONS** (postage extra) ..... $0.158/fiche

**CLEARINGHOUSE MICROFICHE COLLECTIONS** (postage extra) ..... $0.162/fiche

**SPECIAL COLLECTIONS** (postage extra)

| | |
|---|---|
| *Office of Education Research Reports 1956-65* | $ 404.43 |
| *Pacesetters in Innovation, Fiscal Year 1966* | 144.57 |
| *Pacesetters in Innovation, Fiscal Year 1967* | 175.31 |
| *Pacesetters in Innovation, Fiscal Year 1968* | 112.12 |
| *Selected Documents on the Disadvantaged* | 334.28 |
| *Selected Documents in Higher Education* | 153.48 |
| *Manpower Research: Inventory for Fiscal Year 1966 and 1967* | 79.67 |
| *Manpower Research: Inventory for Fiscal Year 1968* | 44.41 |
| *Manpower Research: Inventory for Fiscal Year 1969* | 57.71 |